*Universidade de Brasília*

*Instituto de Ciências Exatas*

*Departamento de Estatística*

*Trabalho de Conclusão de Curso* 2

# MODELOS LINEARES DINÂMICOS APLICADOS EM SÉRIES TEMPORAIS DE CAPTAÇÃO DA CAIXA ECONÔMICA FEDERAL POR POUPANÇA

**Lucas Lopes dos Santos Haine**

**Brasília - DF**

**2014**

**Lucas Lopes dos Santos Haine**

# MODELOS LINEARES DINÂMICOS APLICADOS EM SÉRIES TEMPORAIS DE CAPTAÇÃO DA CAIXA ECONÔMICA FEDERAL POR POUPANÇA

Monografia apresentada para obtenção do título de Bacharel em Estatística.

Orientador: Prof. Dr. Antônio Eduardo Gomes

Brasília - DF
2014

Dedico este trabalho ao meu querido tio Daniel,
que sempre acreditou no meu sucesso.

# Agradecimentos

Primeiramente, agradeço à minha mãe, Arielce Pereira Haine, e ao meu pai, Moacir Lopes dos Santos, que me proporcionaram meios para que eu pudesse ingressar na universidade.

Também sou grato:

ao professor Antônio Eduardo, que possibilitou a realização desta monografia;

à professora Cibele pela disposição em tirar minhas dúvidas;

ao meu grande amigo Adélio, pela ajuda no trabalho, por todos os conselhos e pelos momentos de descontração;

aos demais professores e aos funcionários do departamento de estatística.

# Sumário

# Lista de Figuras

# Resumo

Esta monografia trata do estudo de modelos dinâmicos lineares com aplicação à captação através de poupança da Caixa Econômica Federal. Uma das muitas formas do banco arrecadar dinheiro é através de depósitos em conta poupança. O banco utiliza o dinheiro depositado pelo cliente em diversas operações e, em troca, o cliente recebe um pequeno rendimento mensal sobre o capital depositado.

A caderneta de poupança é um entre diversos produtos oferecidos pela Caixa. Sendo assim, é uma das formas que o banco arrecada dinheiro. Como forma de maximizar o lucro, é necessário que haja planejamento. Esse planejamento se dá na forma de análise de dados temporais, geográficos e outros tipos.

Nesta monografia, será apresentado um método que foca exclusivamente no método quantitativo. Através desse método é possível verificar como se dá o crescimento da caderneta de poupança da Caixa durante o período estudado. Também é possível comparar com outros dentre os principais bancos, para ver como a Caixa figura no mercado nacional.

**Palavras chave:** Modelos Lineares Dinâmicos, Modelos Dinâmicos Lineares, Modelos Dinâmicos, Caixa Econômica Federal, CEF, Caixa, Mercado, séries temporais.

# 1 Introdução

No mercado da atualidade existe uma grande demanda sobre modelos de previsão de séries temporais. Os modelos dinâmicos se apresentam como uma alternativa dentre inúmeros modelos.

É de interesse da Caixa que seja feita uma estimativa dos valores futuros arrecadados, assim como dos parâmetros utilizados no cálculo, para fins de planejamento.

Os nossos dados, a princípio, consistem em uma série temporal de 115 meses, abrangendo dados de captação da Caixa Econômica Federal de janeiro de 2004 a julho de 2013, e da série mercado, de 140 meses, que abrange dados dos quatro maiores bancos do país no período de dezembro de 2001 a julho de 2013. Os bancos em questão são: CEF, Banco do Brasil, Itaú e Bradesco.

A metodologia escolhida para fazer as previsões é a de modelos lineares dinâmicos, que é chamada dessa forma por utilizar parâmetros que podem variar com o tempo. Além da flexibilidade proporcionada pelos parâmetros que variam, esse modelo é apropriado para se trabalhar com séries correlacionadas.

Como os valores são positivos para as duas séries, e o crescimento em um determinado mês é relativo ao mês anterior, faremos nossa análise baseada em uma série de log-retornos dos dados originais. Os log-retornos podem assumir qualquer valor real em teoria, porém na prática sua dispersão costuma ser pequena. Assim sendo, podemos controlar a variância dos dados. Ou seja, uma série de log-retornos é apropriada tanto por se tratar de dados financeiros quanto pela metodologia de modelos dinâmicos.

# 2 Modelos Lineares Dinâmicos

Um modelo dinâmico é representado por um conjunto de quatro matrizes, ou quatro séries de matrizes $\{F_t, G_t, V_t, W_t\}$ que variam com o tempo. Essas matrizes, junto com um vetor $\theta_0$ e uma matriz $C_0$, de parâmetros a priori, determina a forma do modelo.

$$\begin{aligned} Y_t &= F_t'\theta_t + \nu_t, \qquad \nu_t \sim N(0, V_t) \\ \theta_t &= G_t\theta_{t-1} + \omega_t, \qquad \omega_t \sim N(0, W_t) \end{aligned}$$

Esse é o caso geral, onde a cada tempo temos um conjunto de matrizes do modelo, que servem para modelar um vetor de observações $Y_t$. Para maiores detalhes, ver [Petris, Petrone and Campagnoli; 2009] e [WEST, HARRISON; 1997].

Para o nosso estudo, usaremos um caso mais simples, de matrizes constantes. Ou seja, o nosso modelo será representado por um conjunto de matrizes $\{F, G, V, W\}$, onde nossas quatro matrizes são constantes.

$$Y_t = F'\theta_t + \nu_t, \qquad \nu_t \sim N(0,V)$$
$$\theta_t = G\theta_{t-1} + \omega_t, \qquad \omega_t \sim N(0,W)$$

Dependendo da forma como são definidas as matrizes, podemos ter diversos casos distintos de modelos dinâmicos. Alguns exemplos:

Passeio Aleatório:

$$\mu_t = \mu_{t-1} + \omega_t$$

$$y_t = \mu_t + \nu_t$$

Neste caso, as matrizes que definem o modelo serão:

$$F = [1] \qquad G = [1] \qquad \theta_t = \mu_t$$

Crescimento Linear:

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \omega_{t,1}$$

$$\beta_t = \beta_{t-1} + \omega_{t,2}$$

$$y_t = \mu_t + \beta_t \cdot t + \nu_t$$

As matrizes que definem o modelo:

$$\theta_t = \begin{bmatrix} \mu_t \\ \beta_t \end{bmatrix} \qquad G = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad W = \begin{bmatrix} \sigma_\mu^2 & 0 \\ 0 & \sigma_\beta^2 \end{bmatrix} \qquad F = [1 \quad 0]$$

Sazonal (4 períodos):

$$\theta_0 = \begin{bmatrix} s_4 \\ s_3 \\ s_2 \end{bmatrix} \qquad G = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad W = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad F = [1 \quad 0 \quad 0]$$

$$s_1 = -s_2 - s_3 - s_4$$

É possível criar diversos tipos de modelos lineares dinâmicos. Além disso, tipos diferentes de modelos podem ser combinados para criar outro modelo mais geral. Ou seja, podemos ver que os modelos dinâmicos são bastante flexíveis. Usaremos um caso de modelo dinâmico em que combinamos dois modelos distintos no capítulo a seguir.

Para os nossos dados em questão, utilizaremos duas séries de log-retornos. Sendo a matriz conjunta de dados da Caixa Econômica e de Mercado $Y_{140\times 2}$, definimos uma nova matriz $y_{139\times 2}$, onde

$$y_{ij} = ln\left(\frac{Y_{i+1,j}}{Y_{i,j}}\right).$$

A primeira coluna de $y$ representa dados da Caixa, e a segunda representa Mercado. Para cada linha temos observações de um mês, começando no mês de janeiro de 2002, representando o crescimento do Mercado relativo ao mês anterior. Como os dados da Caixa estão indisponíveis para meses anteriores a janeiro de 2004, só temos log-retornos da Caixa a partir de fevereiro de 2004. Ou seja, os valores de $y_{1\,1}$ a $y_{1\,25}$ são desconhecidos, e portanto não são utilizados no nosso modelo.

# 3 Caixa

Neste estudo, usaremos um modelo de tendência linear com sazonalidade. Nossa metodologia nos permite obter tal modelo somando-se dois modelos mais simples, separando tendência e sazonalidade.

Para o modelo de tendência linear, utilizamos um modelo polinomial de ordem dois (a saber, uma reta), com os parâmetros iniciais iguais aos valores estimados por um modelo de regressão.

Já para o modelo sazonal, faremos uma sazonalidade de 12 períodos, utilizando os valores fornecidos pela série nos meses de março de 2004 a janeiro de 2005. Com isso, estamos assumindo que a sazonalidade é constante, e não fornecemos o valor sazonal para fevereiro, que seria o primeiro mês disponível, pois a soma dos 12 valores sazonais tem 11 graus de liberdade. Os parâmetros sazonais são calculados da seguinte forma:

$$s = \begin{bmatrix} s_{12} \\ s_{11} \\ s_{10} \\ s_9 \\ s_8 \\ s_7 \\ s_6 \\ s_5 \\ s_4 \\ s_3 \\ s_2 \end{bmatrix}_{11\times 1},$$

onde $s_i = y_{i+25,1} - \frac{1}{11}\sum_{k=27}^{37} y_{k1}$, $i = 2, \ldots, 12$, e $s_1$, que não aparece explicitamente no modelo, é $s_1 = -\sum_{k=2}^{12} s_k$. Neste caso, $s_1$ representa a sazonalidade de fevereiro de 2004, que é o primeiro mês para o qual temos um log-retorno disponível. Os outros parâmetros representam a sazonalidade tendo por base os meses seguintes, até janeiro de 2005. Após esse período, os valores se repetem.

Os parâmetros para o modelo são os seguintes:

$$F = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{13\times 1} \qquad V = \left[\frac{1}{20.000}\right]$$

$$G = \begin{bmatrix} 1 & 1 & & & 0 & \\ 0 & 1 & & & & \\ & & -1 & \dots & -1 & -1 \\ & 0 & 1 & 0 & 0 & 0 \\ & & 0 & \ddots & 0 & \vdots \\ & & 0 & 0 & 1 & 0 \end{bmatrix}_{13\times 13}$$

$$W = W_{13\times 13}, \qquad \text{onde} \begin{cases} w_{22} = w_{33} = \dfrac{1}{300.000} \\ w_{ij} = 0 \ c.c. \end{cases}$$

$$\theta_0 = \begin{bmatrix} 0{,}00847 \\ 0{,}00050 \\ -0{,}00055 \\ 0{,}01443 \\ -0{,}00394 \\ -0{,}00825 \\ -0{,}00848 \\ -0{,}00528 \\ 0{,}00850 \\ 0{,}00261 \\ 0{,}01394 \\ -0{,}00598 \\ -0{,}00702 \end{bmatrix}_{13\times 1}$$

$$C_0 = 6{,}140351 \cdot 10^{-7} \cdot I_{13\times 13}$$

A matriz de delineamento é dada dessa forma porque temos duas parcelas de $\theta_t$ que entram no cálculo de $y_t$. São elas a primeira, representando o modelo linear, e a terceira, que representa o parâmetro sazonal. A matriz de evolução pode ser separada em duas matrizes quadradas, onde a primeira representa o modelo polinomial e a segunda representa o modelo sazonal. Os valores restantes são nulos, indicando que as componentes linear e sazonal são independentes entre si.

A razão entre as variâncias determina se o modelo se aproxima mais de uma tendência geral ou dos dados observados. Valores mais altos de $V$ indicam que as estimativas se aproximam de uma tendência dada simplesmente pelo modelo, enquanto que valores mais altos de $W$ indicam que as estimativas são mais próximas dos dados observados.

Um dos motivos para a escolha dos valores acima foi para que as estimativas do modelo alisado tivessem um equilíbrio entre a informação a priori e a observada através dos dados ao longo do tempo. Outro critério para a escolha das variâncias, que reflete nos valores pequenos, foi para que a proporção dos valores dentro do intervalo de confiança se aproximasse da proporção esperada.

Também podemos usar a matriz de variância de evolução para trabalhar com fator de desconto. O fator de desconto é uma medida do quanto a informação do mês anterior é agregada ao mês em questão. Neste modelo, e em todos os outros que seguem, não foi usado fator de desconto por uma questão de simplicidade. Para maiores informações, o leitor pode consultar [WEST, HARRISON; 1997] e [Petris, Petrone and Campagnoli; 2009], e para uma abordagem mais simples, convém consultar [Da Silva Neto; 2010].

O vetor $\theta_0$ fornece dados da tendência linear, que representam o valor esperado no mês inicial e crescimento esperado a cada mês, e valores sazonais, que se alternam periodicamente, em um período de 12 meses.

A matriz $C_0$ representa a variância a priori dos dados. Enquanto essa variância não chega a afetar diretamente as estimativas pontuais, ela altera o intervalo de confiança para as previsões e também tem grande relevância para simulações.

Nosso modelo não detecta quebras no comportamento da série, que podem ser causadas por crises financeiras ou outros fatores externos. Contudo, ele serve para que tenhamos uma ideia geral acerca do crescimento esperado.

Logo abaixo, vemos uma série de previsões, utilizando apenas os parâmetros iniciais:

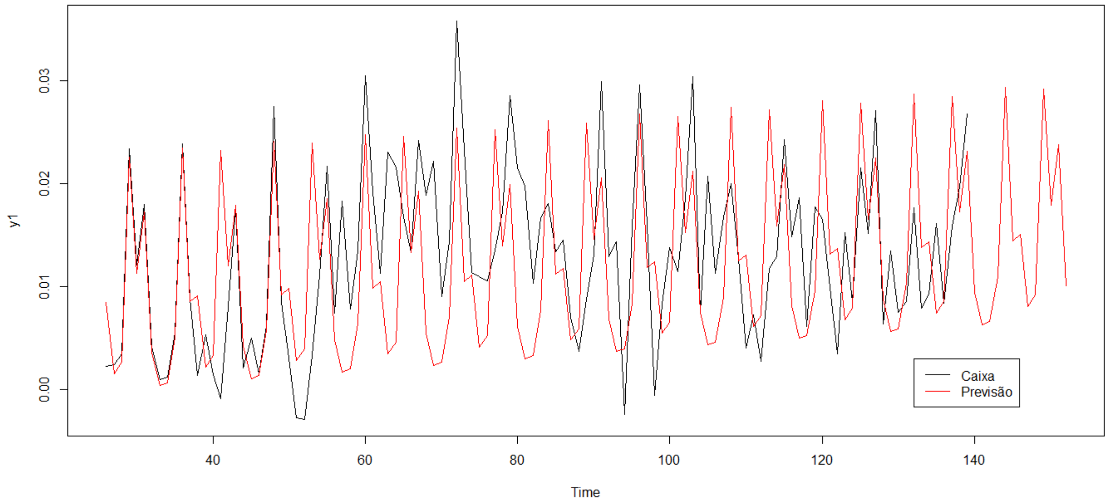

Figura 1: Modelo Preditivo - Caixa

A curva preta representa os dados observados, enquanto a vermelha representa a previsão. Podemos ver claramente que há um crescimento ao longo do tempo, assim como flutuações sazonais, de acordo com o modelo que selecionamos. A previsão para os 12 primeiros meses é muito próxima devido à forma como escolhemos nossos parâmetros sazonais. A previsão pode ser estendida o quanto quisermos, porém os valores estimados perdem consistência para tempos mais distantes.

Apesar de ser uma boa representação, essa série não agrega a informação dada pelos dados observados ao longo da série. E se usássemos esse mesmo modelo, porém usando todos os valores observados até o tempo de uma observação específica?

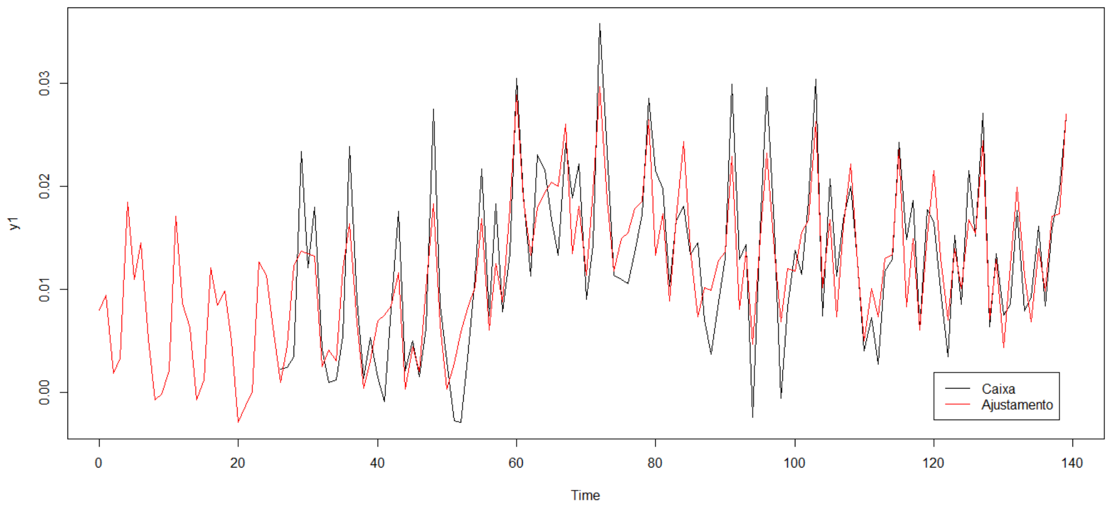

Figura 2: Modelo Alisado - Caixa

Agora os valores estimados não se aproximam tanto dos dados dos primeiros meses, mas a informação agregada gera estimativas com viés pequeno para os dados observados mais à frente. Perceba que temos estimativas até mesmo para os meses não disponíveis. O primeiro valor é uma média para o mês inicial, usando os parâmetros iniciais do modelo, e em seguida vemos um decréscimo na tendência até o primeiro mês observado, onde a estimativa começa a se adequar aos dados.

Vimos que as estimativas para os valores da série são pouco enviesados. Mas é de nosso interesse ver se eles de fato são consistentes. Através da função dlmSmooth do pacote dlm, e usando uma álgebra matricial, podemos calcular a variância, e consequentemente a distribuição, dos nossos estimadores. O cálculo é feito da seguinte forma:

$$var(\hat{y}_t) = F_t' \cdot U_t \cdot D_t^2 \cdot U_t' \cdot F_t,$$

onde $U_{n\times n}$ é uma matriz ortogonal e $D_{n\times n}$ é uma matriz diagonal, ambas dadas pelo método SVD.

Sabendo que nossos estimadores têm distribuição normal, com média $F_t' \cdot \theta_t$ e variância dada pelo cálculo acima, podemos criar um intervalo de credibilidade. Ele servirá para vermos o quanto as nossas estimativas de fato são consistentes.

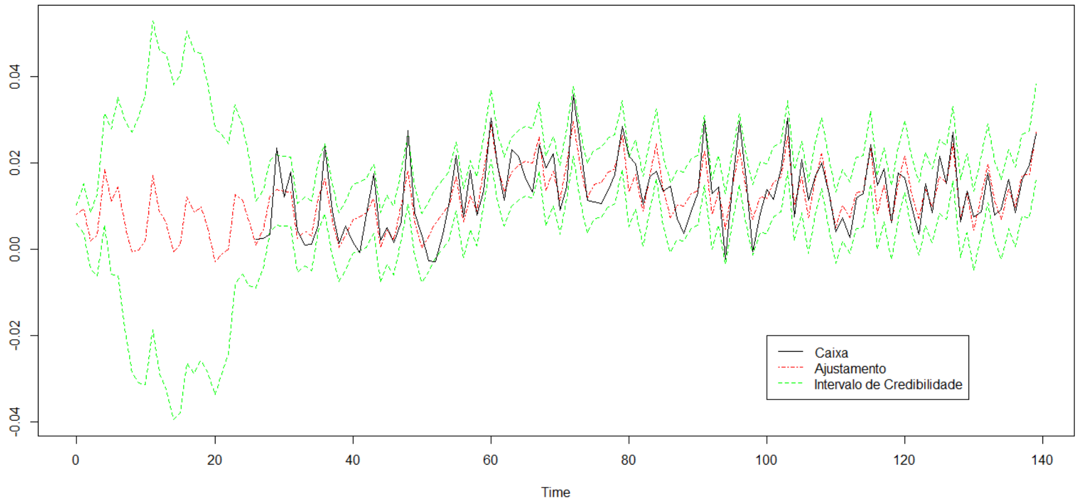

Figura 3: Modelo Ajustado para CEF com Intervalo de 95% de Credibilidade

O gráfico acima mostra um intervalo com 95% de credibilidade. Não temos uma visão clara de crescimento da série. Porém, podemos notar que para a maioria dos meses, o valor esperado é maior que 0. Ou seja, em condições normais, espera-se sempre um crescimento da captação por poupança da Caixa.

# 4 Mercado

Agora iremos modelar a série de mercado, utilizando o mesmo método empregado para os dados da Caixa Econômica Federal. Esta modelagem visa comparar o crescimento da Caixa com o restante do mercado. Para esta série, temos dados temos mais dados em meses anteriores aos da série anterior, totalizando 139 meses. Os log-retornos para esta série estão disponíveis para um período de janeiro de 2002 a julho de 2013, como foi definido na seção 2. Poderíamos modelar as duas séries conjuntamente, mas este modelo foi usado em detrimento de um caso mais geral por questão de simplicidade.

Para calcular os parâmetros sazonais iniciais, utilizei os dados dos meses de fevereiro a dezembro de 2004. Decidi utilizar os dados desse ano em detrimento dos dados de 2002, que apresentam uma irregularidade. Se dispuséssemos de mais conhecimento a priori, os dados certamente seriam distantes do que observamos em 2002, por se tratar de um ano com grandes flutuações. Porém, esse tipo de informação requer uma análise econômica mais aprofundada, que foge ao escopo deste trabalho.

As variâncias de evolução e do sistema novamente foram escolhidas de forma que houvesse um equilíbrio entre a informação a priori e a observada através dos dados, e também produzir intervalos de confiança coerentes com os dados.

Os parâmetros do modelo seguem logo abaixo:

$$\theta_0 = \begin{bmatrix} 0{,}00675 \\ 0{,}00006 \\ 0{,}01980 \\ 0{,}00093 \\ -0{,}00357 \\ -0{,}00206 \\ -0{,}00698 \\ 0{,}00534 \\ 0{,}00417 \\ 0{,}00852 \\ -0{,}00580 \\ -0{,}01345 \\ -0{,}00691 \end{bmatrix}_{13\times 1}$$

$$C_0 = 8{,}23741 \cdot 10^{-7} \cdot I_{13\times 13}$$

$$V = \left[\frac{1}{12000}\right]$$

$$W = W_{13\times 13}, \qquad \text{onde} \begin{cases} w_{22} = w_{33} = \frac{1}{240.000} \\ w_{ij} = 0 \; c.c. \end{cases}$$

Todas demais matrizes são iguais às utilizadas no modelo da Caixa.

Agora veremos uma previsão partindo do mês inicial, nos mesmos moldes do realizado anteriormente.

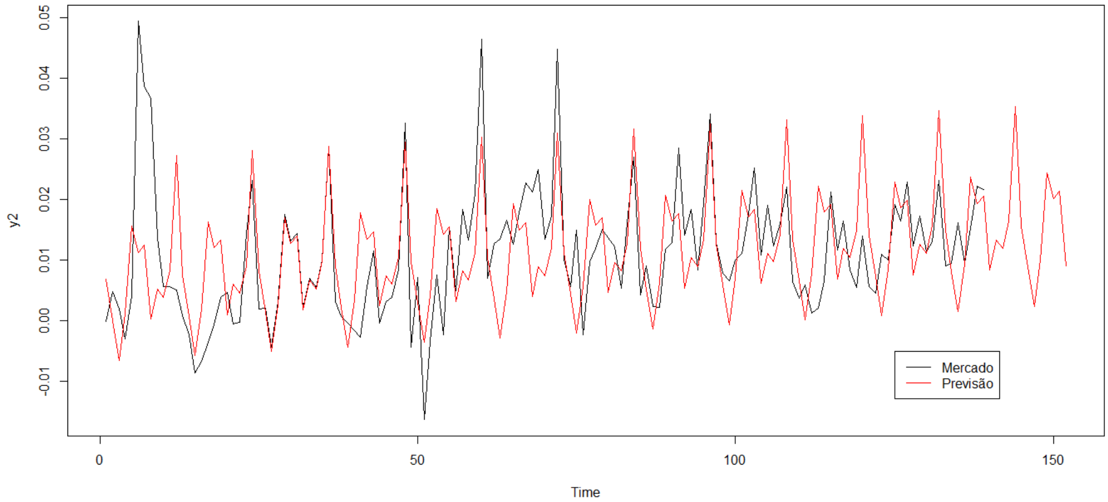


Figura 4: Modelo Preditivo - Mercado

Novamente, temos uma série de previsões partindo do mês inicial, que não agregam novas informações. Mesmo assim, em alguns pontos ao longo da série os resultados são muito próximos.

Tendo visto um modelo preditivo partindo dos parâmetros a priori, também é interessante olhar o modelo alisado, com estimativas que utilizam os dados observados. Segue um gráfico de tal modelo logo abaixo:

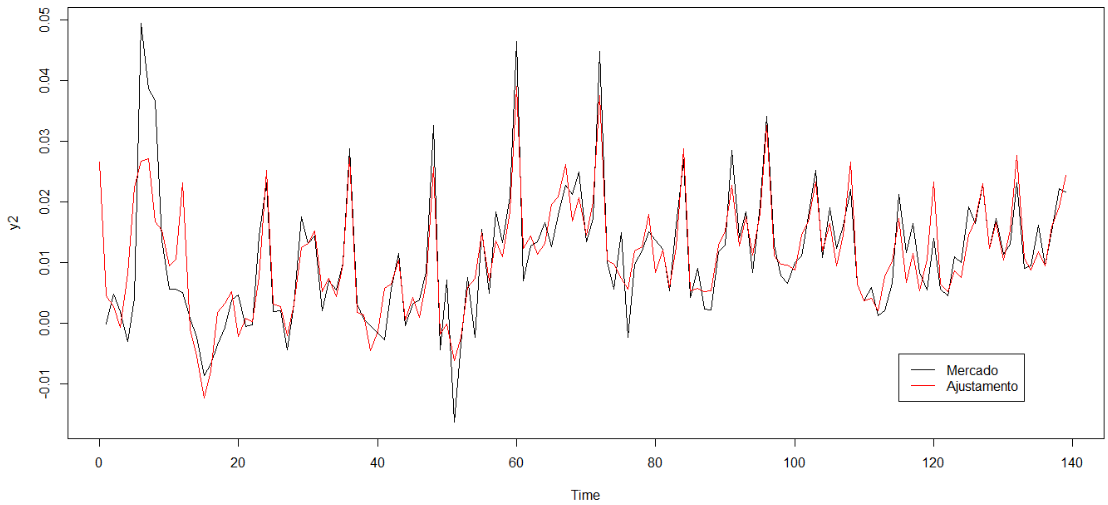


Figura 5: Modelo Alisado - Mercado

Podemos notar que as estimativas são consistentes a partir do ano de 2004. Para um modelo de média simples, não é esperado uma aproximação tão precisa dos dados. Contudo, como assumimos uma tendência de crescimento linear e variações sazonais a priori, é de se esperar que os dados estimados estejam próximos dos reais.

Existe uma mudança de comportamento com grandes variações no meio do período observado. Porém, próximo ao fim, a série se torna mais comportada, com uma tendência mais constante e variações menores.

É de nosso interesse que tenhamos um intervalo de credibilidade para as estimativas do mercado. Através do mesmo método empregado para o modelo da Caixa, podemos analisar o modelo de Mercado.

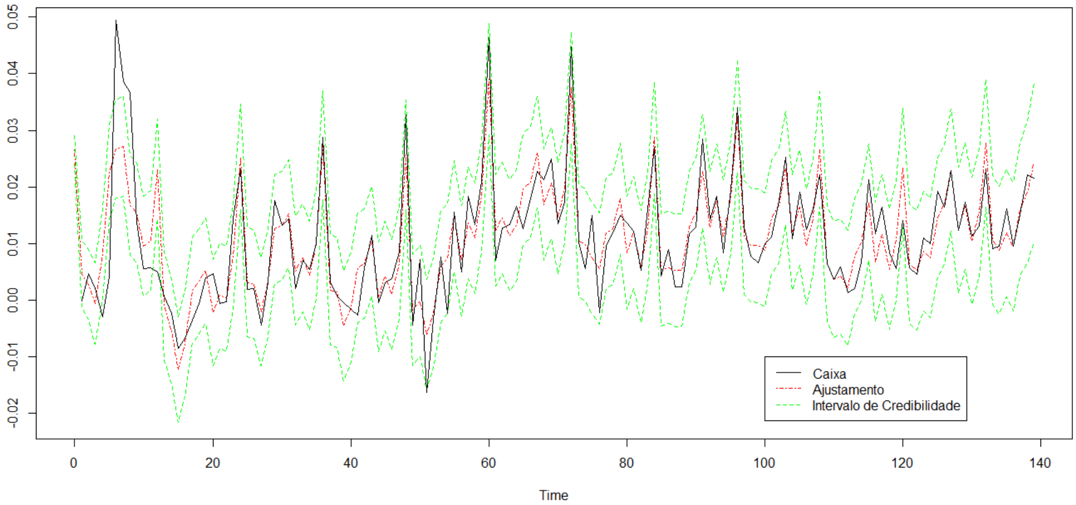


Figura 6: Modelo Ajustado para Mercado com Intervalo de 95% de Credibilidade

Assim como na Caixa, observamos muitos valores positivos. Porém, esta série apresenta mais irregularidades. Por se tratar de uma série que engloba dados dos quatro maiores bancos brasileiros nesse período, esta série está mais sujeita a variações devido a crises, ciclos econômicos e outros fatores externos. Mesmo com mais dados disponíveis, as estimativas deste modelo são menos consistentes que as da Caixa.

# 5 Comparação entre Caixa e Mercado

Os dados observados, tanto para a série da Caixa quanto para Mercado são dados de log-retorno. Isso significa que cada valor representa o quanto a série cresceu em relação ao mês anterior. Sendo assim, mesmo que a variável mercado dos dados originais contenha os valores da Caixa, entre outros bancos, o log-retorno dessa série não necessariamente será maior se comparado ao log retorno da Caixa unicamente. A série Mercado serve para que analisemos se a Caixa Econômica Federal de fato está fazendo frente aos outros bancos no quesito poupança.

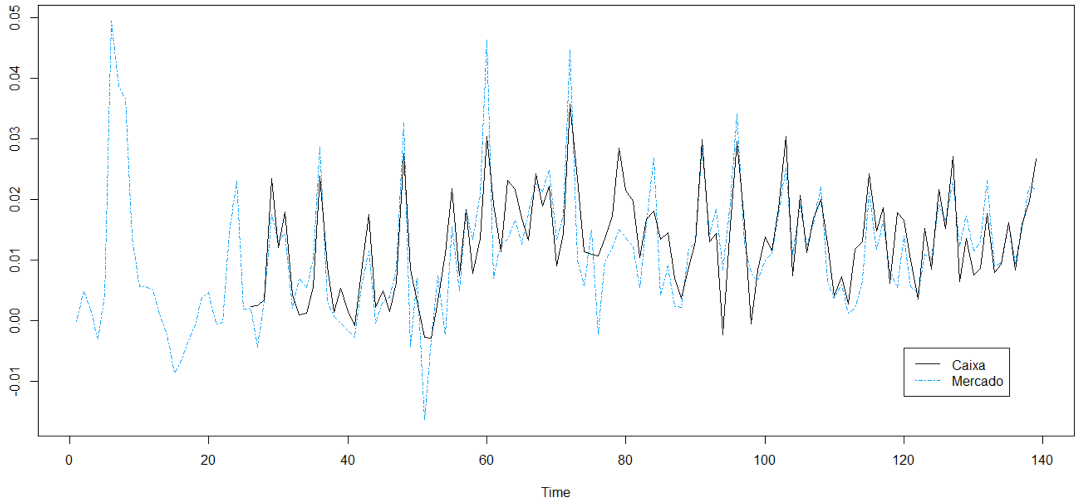

Figura 7: Caixa & Mercado

O gráfico acima relaciona as duas séries. As flutuações do mercado são maiores, mas isso é a única coisa que podemos inferir do gráfico. Apesar do crescimento do mercado ser maior em alguns meses, em outros observamos um crescimento relativo maior da Caixa. De fato, se formos olhar o somatório de cada série, temos:

$$\sum_{i=26}^{139} y_{i1} = 1{,}478727$$

$$\sum_{i=26}^{139} y_{i2} = 1{,}345184$$

Isso significa que, durante o período que comporta os últimos 114 meses observados, a caderneta de poupança da Caixa Econômica cresceu em 4,39 vezes, ao passo que o crescimento para os bancos em geral foi de 3,84 vezes.

Apesar dessa diferença, em alguns pontos as duas séries são bem próximas. Será que não podemos utilizar o modelo empregado para o mercado para modelar a série de log-retornos da Caixa?

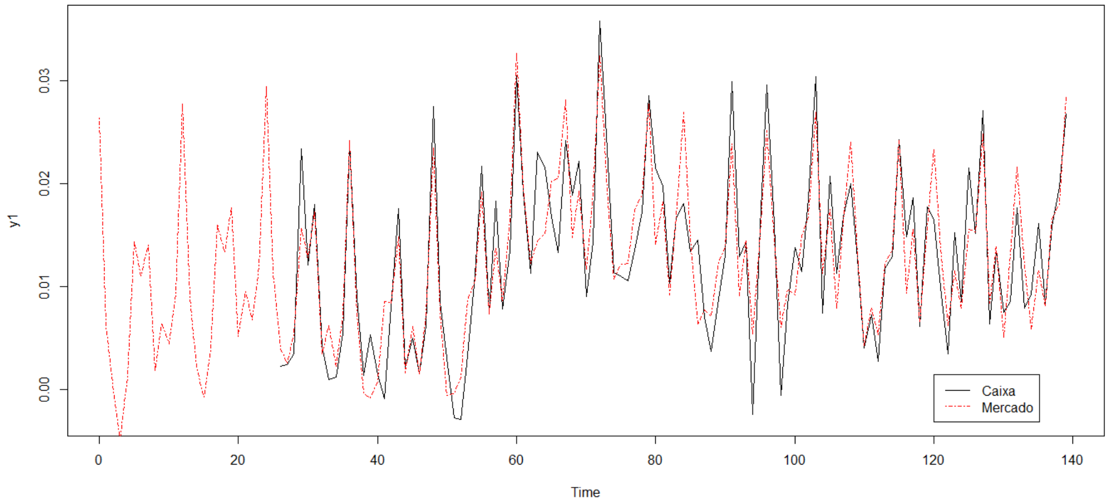


Figura 8: Modelo Alisado para Caixa com parâmetros de Mercado

Este modelo tem os parâmetros usados para o mercado a priori, mas está ajustado para os dados da Caixa Econômica. Fazendo o intervalo de credibilidade, podemos ver se de fato o modelo é consistente para esses dados.

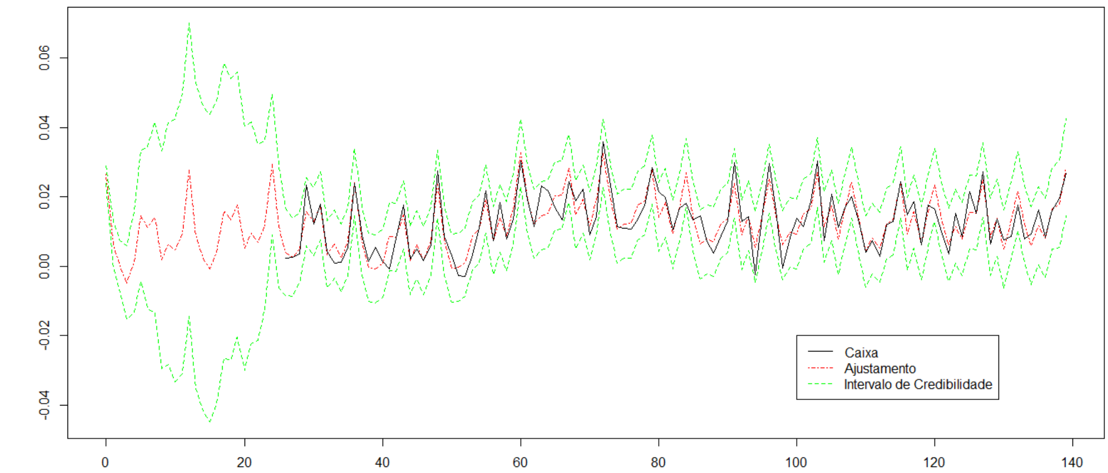


Figura 9: Modelo Alisado para Caixa com parâmetros de Mercado, com intervalo de 95% de credibilidade

Esse modelo não difere tanto do primeiro modelo para os dados da Caixa. A priori, tanto a Caixa quanto o Mercado têm a mesma informação, exceto para períodos de crise financeira. Podemos então, com base nos dados observados durante esse tempo, dizer que o crescimento da Caixa não está abaixo do crescimento médio do mercado.

# 6 Outros Modelos

Neste capítulo, veremos se o modelo que utilizamos é mais apropriado que outros modelos usados em séries temporais. O nosso modelo pode ser comparado com qualquer outro modelo para séries temporais. Entre eles, temos o modelo de regressão, que é o mais simples, diversos casos particulares e inclusive alguns mais gerais de modelos ARIMA (e.g., auto regressivo, médias móveis, ARMA). Também podemos usar diversos modelos de alisamento exponencial, entre eles, o de Holt-Winters. Porém, este último não será utilizado devido à dificuldade de se analisar pelo método em questão, além de já ser bastante semelhante ao modelo que temos em mãos.

Neste estudo, usaremos o critério de erro quadrático para previsão um passo a frente. A estatística de erro quadrático será definida da seguinte forma:

$$EQ = \sum_{i=26}^{139} (\hat{y}_{i1} - y_{i1})^2,$$

onde $y_{i1}$ é o log-retorno da Caixa observado no i-ésimo mês e $\hat{y}_{i1}$ é a previsão um passo a frente estimada no mesmo mês em questão.

Os modelos utilizados são: um modelo ARMA(2,2), ARIMA(5,1,2), SARIMA(3,1,2)(2,0,2), SARIMA(1,1,3)(1,0,1), o modelo dinâmico que combina tendência linear crescente e variação sazonal – o principal do nosso estudo –, e outro modelo dinâmico que utiliza parâmetros a priori com base em dados do mercado.

Já vimos anteriormente o comportamento dos dois modelos dinâmicos em questão. Contudo, também é interessante vermos como se comportam os outros modelos aos quais queremos comparar:

| Modelo | EQ |
|---|---|
| DLM CEF | 0,00189 |
| DLM CEF/MER | 0,00146 |
| ARMA(2,2) | 0,00597 |
| ARIMA(5,1,2) | 0,00623 |
| SARIMA(3,1,2)(2,0,2) | 0,00325 |
| SARIMA(1,1,3)(1,0,1) | 0,00322 |

Tabela 1: Comparação de modelos pelo EQ

Modelos com uma soma de erros quadráticos menor são mais precisos. Os dois primeiros modelos são os descritos nas seções 3 e 5, respectivamente. Todos os outros modelos tiveram seus parâmetros e previsões estimados usando métodos da estatística frequentista.

Como podemos notar, os modelos dinâmicos são muito mais precisos, e em especial, o segundo modelo nos dá a melhor precisão. Podemos associar isso à informação a priori que utilizamos no nosso estudo. Como já foi visto anteriormente, ela descreve os dados com bastante precisão. Caso os parâmetros do segundo modelo não estivessem disponíveis, o primeiro modelo poderia ser usado sem grande perda de informação.

Os dois modelos SARIMA são os mais precisos dentre os modelos frequentistas. A diferença na estatística EQ chega a ser insignificante, mas o SARIMA(1,1,3)(1,0,1) ainda é preferível por uma questão de simplicidade.

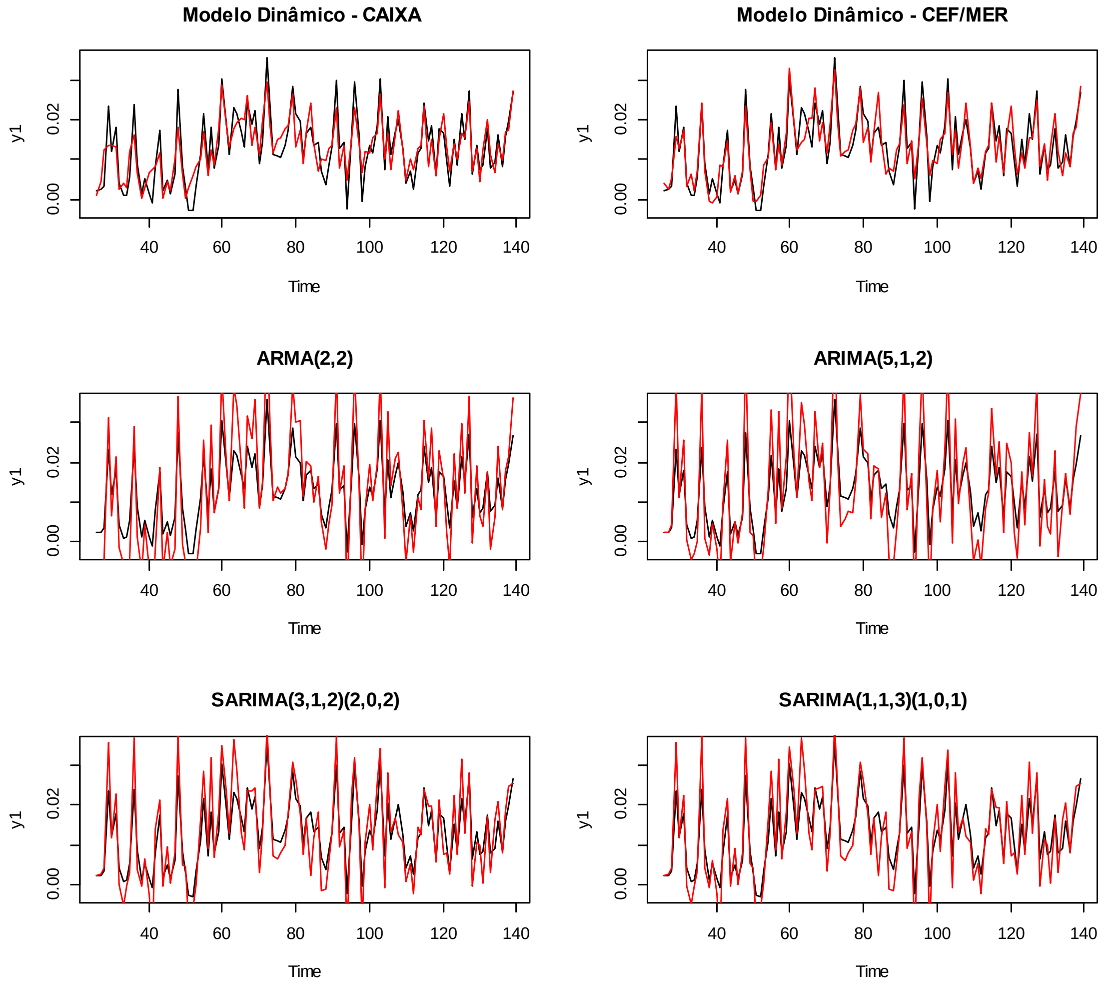


Figura 10: Comparação de Modelos

# 7 Conclusão

Durante o período estudado, a caderneta de poupança da Caixa Econômica Federal cresceu consideravelmente. E pelo que mostram os dados, a tendência é que cresça ainda mais. Porém, devemos tomar cuidado ao fazer inferência para valores futuros, pois estes perdem consistência. Principalmente em se tratando de dados econômicos, pois a qualquer momento uma crise pode mudar completamente o quadro observado.

Dada a grande precisão do nosso modelo, podemos assumir que a metodologia de modelos lineares dinâmicos é apropriada para analisar os dados em questão. O modelo utilizado mostrou-se superior até mesmo ao ARMA, que é um modelo tradicionalmente usado em dados econômicos de log retornos. Possivelmente existem outros modelos mais gerais, bayesianos ou clássicos, que descrevem os dados de forma mais precisa. Porém, uma vez entendida a metodologia em questão, o nosso modelo pode ser facilmente aplicado.

Apesar de haver mais resultados que poderiam ser implementados, o resultado foi satisfatório, pois de fato foi possível analisar os dados usando o modelo proposto, com parâmetros que são de interesse da Caixa. Com isso, espero acrescentar aos trabalhos de modelos dinâmicos com aplicações reais e que esta monografia sirva de referência para trabalhos futuros.

# Referências Bibliográficas


- Petris, Petrone, and Campagnoli. Dynamic Linear Models with R. Springer (2009).
- WEST, Mike; HARRISON, Jeff. Bayesian Forecasting and Dynamic Models. Second Edition. Springer (1997).
- R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
- Giovanni Petris (2010). An R Package for Dynamic Linear Models. Journal of Statistical Software, 36(12), 1-16. URL http://www.jstatsoft.org/v36/i12/.
- Da Silva Neto, Adélio Henrique. Modelos Dinâmicos Lineares. 2010.